

Jihao Xin

✉ jihao.xin@kaust.edu.sa
🌐 <https://jihaoxin.github.io/>

in jihao-xin
🎂 Oct 27, 1998

🇨🇳 China



Introduction

I'm a PhD student in computer science, working with Prof. Marco Canini at SANDS Lab. My research focuses on accelerating large scale AI through algorithm-system co-designed approaches. I'm recently working on LLM compression and GPU communication.

Education

- | | |
|-----------------------------|---|
| January 2023 – Present | 📖 Ph.D. student in Computer Science , KAUST, Saudi Arabia. |
| August 2021 – December 2022 | 📖 Master of Science in Computer Science , KAUST, Saudi Arabia.
<i>Dean's List Award</i>
GPA = 3.88/4.0 |
| October 2020 – October 2021 | 📖 Master of Science in Applied Computational Science and Engineering , Imperial College London, UK.
<i>Distinction Degree</i> |
| September 2016 – June 2020 | 📖 Bachelor of Engineering in Computer Science and Technology , Shandong University, China.
GPA = 89.34/100 |

Employment & Internship

- | | |
|-----------------------------|--|
| September 2023 – March 2024 | 📖 Research Intern , Microsoft Research Asia, Beijing, China.
Supervised by Dr. Changho Hwang
Efficient GPU communication for distributed AI tasks. |
| May 2022 – August 2022 | 📖 Research Intern , ISTAustria, Klosterneuburg, Austria.
Supervised by Prof. Dan Alistarh
TopK compression with efficient CUDA-based threshold selection. |
| August 2019 – November 2019 | 📖 Engineering Intern , Tencent Cloud, Shenzhen, China.
Tencent international website backend development by GoLang. |
| June 2019 – August 2019 | 📖 Research Intern , The University of Western Australia, Remote.
YOLO-based splash detection of Western Australia's coastal highway by Python. |

Research Publications

- 1 J. Xin, M. Canini, P. Richtárik, and S. Horvath, "Quantize once, train fast: Allreduce-compatible compression with provable guarantees," in *ECAI*, Bologna, Italy, 2025.
- 2 J. Weng, B. Han, D. Gao, R. Gao, W. Zhang, A. Zhong, C. Xu, J. Xin, Y. Luo, L. W. Wills, and M. Canini, "Assassyn: A unified abstraction for architectural simulation and implementation," in *ISCA*, Tokyo, Japan, 2025.
- 3 N. Xue, D. Guo, J. Zhang, J. Xin, Z. Li, and X. Huang, "Openfunction for software defined iot," in *ISNCC*, **Best Paper Award**, Dubai, UAE, 2021.

- 4 **J. Xin**, M. Canini, P. Richtárik, and S. Horváth, “Global-qsgd: Allreduce-compatible quantization for distributed learning with theoretical guarantees,” in *EuroMLSys*, Rotterdam, Netherlands, 2025.
- 5 **J. Xin**, S. Bae, K. Park, M. Canini, and C. Hwang, “Immediate communication for distributed ai tasks,” in *HotInfra*, Austin, USA, 2024.
- 6 **J. Xin**, I. Ilin, S. Zhang, M. Canini, and P. Richtárik, “Kimad: Adaptive gradient compression with bandwidth awareness,” in *DistML*, Paris, France, 2023.
- 7 M. Mao, X. Feng, **J. Xin**, and T. W. S. Chow, “A convolutional neural network-based maximum power point voltage forecasting method for pavement pv array,” *IEEE TIM*, 2023.

Skills

Programming	■ C++/CUDA, Python/PyTorch, GoLang
Systems	■ Linux, GPU Programming, MPI, NCCL
Languages	■ English (Professional), Chinese (Native)

Miscellaneous Experience

Teaching Asistant

2025 Spring	■ KAUST CS398 Graduate Seminar
2022 Fall, 2024 Fall	■ KAUST CS240: Computing Systems and Concurrency
2023 Spring	■ Saudi Ministry of Interior Training Program CS229: Machine Learning
2022 Summer, 2023 Summer	■ UESTC&KAUST Summer Camp

Seasonal Programs

2019 Summer	■ Summer Research Program , The University of Western Australia.
2019 Winter	■ Robotics & AI Winter School , Imperial College London Hamlyn Center.

Leadership

2024-2025	■ KAUST Graduate Student Council Representative (12/1683).
-----------	---